Chapter 1

Lorentz Group and Lorentz Invariance

In studying Lorentz-invariant wave equations, it is essential that we put our understanding of the Lorentz group on firm ground. We first define the Lorentz transformation as any transformation that keeps the 4-vector inner product invariant, and proceed to classify such transformations according to the determinant of the transformation matrix and the sign of the time component. We then introduce the generators of the Lorentz group by which any Lorentz transformation continuously connected to the identity can be written in an exponential form. The generators of the Lorentz group will later play a critical role in finding the transformation property of the Dirac spinors.

1.1 Lorentz Boost

Throughout this book, we will use a unit system in which the speed of light c is unity. This may be accomplished for example by taking the unit of time to be one second and that of length to be $2.99792458 \times 10^{10}$ cm (this number is *exact*[1]), or taking the unit of length to be 1 cm and that of time to be $(2.99792458 \times 10^{10})^{-1}$ second. How it is accomplished is irrelevant at this point.

Suppose an inertial frame K (space-time coordinates labeled by t, x, y, z) is moving with velocity β in another inertial frame K' (space-time coordinates labeled by t', x', y', z') as shown in Figure 1.1. The 3-component velocity of the origin of K

[1] One cm is *defined* (1983) such that the speed of light in vacuum is $2.99792458 \times 10^{10}$ cm per second, where one second is defined (1967) to be 9192631770 times the frequency of the hyper-fine splitting of the Cs^{133} ground state.

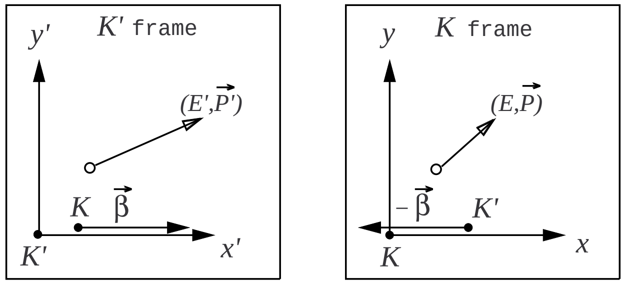

Figure 1.1: The origin of frame K is moving with velocity $\vec{\beta} = (\beta, 0, 0)$ in frame K', and the origin of frame K' is moving with velocity $-\vec{\beta}$ in frame K. The axes x and x' are parallel in both frames, and similarly for y and z axes. A particle has energy momentum $(E, \vec{P})$ in frame K and $(E', \vec{P}')$ in frame K'.

measured in the frame K', $\vec{\beta}'_K$, is taken to be in the $+x'$ direction; namely,

$$\vec{\beta}'_K \text{ (velocity of } K \text{ in } K') = (\beta, 0, 0) \stackrel{\text{def}}{\equiv} \vec{\beta}. \tag{1.1}$$

Assume that, in the frame K', the axes x, y, z are parallel to the axes x', y', z'. Then, the velocity of the origin of K' in K, $\vec{\beta}_{K'}$, is

$$\vec{\beta}_{K'} = -\vec{\beta}'_K = (-\beta, 0, 0) \quad \text{(velocity of } K' \text{ in } K). \tag{1.2}$$

Note that $\vec{\beta}'_K$ ($\vec{\beta}_{K'}$) is measured with respect to the axes of K' (K).

If a particle (or any system) has energy and momentum $(E, \vec{P})$ in the frame K, then the energy and momentum $(E', \vec{P}')$ of the same particle viewed in the frame K' are given by

$$\begin{aligned} E' &= \frac{E + \beta P_x}{\sqrt{1-\beta^2}}, & P'_y &= P_y, \\ P'_x &= \frac{\beta E + P_x}{\sqrt{1-\beta^2}}, & P'_z &= P_z,. \end{aligned} \tag{1.3}$$

This can be written in a matrix form as

$$\begin{pmatrix} E' \\ P'_x \end{pmatrix} = \begin{pmatrix} \gamma & \eta \\ \eta & \gamma \end{pmatrix} \begin{pmatrix} E \\ P_x \end{pmatrix}, \qquad \begin{pmatrix} P'_y \\ P'_z \end{pmatrix} = \begin{pmatrix} P_y \\ P_z \end{pmatrix} \tag{1.4}$$

with

$$\gamma \equiv \frac{1}{\sqrt{1-\beta^2}}, \qquad \eta \equiv \beta\gamma = \frac{\beta}{\sqrt{1-\beta^2}}. \tag{1.5}$$

Note that γ and η are related by

$$\boxed{\gamma^2 - \eta^2 = 1}. \tag{1.6}$$

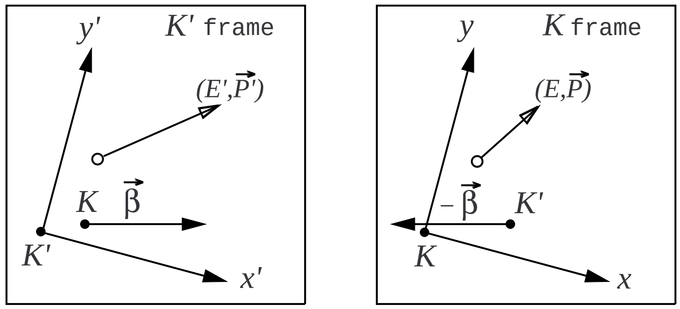

Figure 1.2: Starting from the configuration of Figure 1.1, the same rotation is applied to the axes in each frame. The resulting transformation represents a general Lorentz boost.

Now start from Figure 1.1 and apply the *same rotation* to the axes of K and K' *within each frame* (Figure 1.2). Suppose the rotation is represented by a 3×3 matrix R. Then, the velocity of K' in K, $\vec\beta_{K'}$, and and the velocity of K in K', $\vec\beta'_K$, are rotated by the same matrix R,

$$\vec\beta'_K \to R\vec\beta'_K\,, \qquad \vec\beta_{K'} \to R\vec\beta_{K'}\,, \tag{1.7}$$

and thus we still have

$$\vec\beta'_K = -\vec\beta_{K'} \overset{\text{def}}{\equiv} \vec\beta\,, \tag{1.8}$$

where we have also redefined the vector $\vec\beta$ which is well-defined in both K and K' frames in terms of $\vec\beta_{K'}$ and $\vec\beta'_K$, respectively. The transformation in this case can be obtained by noting that, in (1.4), the component of momentum transverse to $\vec\beta$ does not change and that P_x, P'_x are the components of $\vec P$, $\vec P'$ along $\vec\beta$ in each frame. Namely, the tranformation can be written as

$$\boxed{\begin{pmatrix} E' \\ P'_\parallel \end{pmatrix} = \begin{pmatrix} \gamma & \eta \\ \eta & \gamma \end{pmatrix} \begin{pmatrix} E \\ P_\parallel \end{pmatrix}\,, \qquad \vec P'_\perp = \vec P_\perp}\,, \tag{1.9}$$

where $_\parallel$ and $_\perp$ denote components parallel and perpecdicular to $\vec\beta$, repectively. Note that $\vec P'_\perp$ and $\vec P_\perp$ are 3-component quantities and the relation $\vec P'_\perp = \vec P_\perp$ holds component by component because we have applied the same roation R in each frame.

The axes of K viewed in the frame K' are no longer perpendicular to each other since they are contracted in the direction of $\vec\beta'_K$. Thus, the axes of K in general are not parallel to the corresponding axes of K' at any time. However, since the same rotation is applied in each frame, and since components transverse to $\vec\beta$ are the same in both frames, the corresponding axes of K and K' are exactly parallel when projected onto a plane perpendicular to $\vec\beta$ in either frames. The transformation (1.9) is thus correct for the specific relative oprientation of two frames as defined here, and such transformation is called a Lorentz boost, which is a special case of *Lorentz transformation* defined later in this chapter for which the relative orientation of the two frames is arbitrary.

1.2 4-vectors and the metric tensor $g_{\mu\nu}$

The quantity $E^2 - \vec{P}^2$ is invariant under the Lorentz boost (1.9); namely, it has the same numerical value in K and K':

$$\begin{aligned} E'^2 - \vec{P}'^2 &= E'^2 - (P'^2_\parallel + \vec{P}'^2_\perp) \\ &= (\gamma E + \eta P_\parallel)^2 - \left[(\eta E + \gamma P_\parallel)^2 + \vec{P}^2_\perp\right] \\ &= \underbrace{(\gamma^2 - \eta^2)}_{1} E^2 + \underbrace{(\eta^2 - \gamma^2)}_{-1} P^2_\parallel - \vec{P}^2_\perp \\ &= E^2 - \vec{P}^2 , \end{aligned} \tag{1.10}$$

which is the invariant mass squared m^2 of the system. This invariance applies to any number of particles or any object as long as E and $\vec{P}$ refer to the same object.

The relative minus sign between E^2 and $\vec{P}^2$ above can be treated elegantly as follows. Define a 4-vector P^μ ($\mu = 0, 1, 2, 3$) by

$$P^\mu = (P^0, P^1, P^3, P^4) \stackrel{\text{def}}{\equiv} (E, P_x, P_y, P_z) = (E, \vec{P}) \tag{1.11}$$

called an energy-momentum 4-vector where the index μ is called the Lorentz index (or the space-time index). The $\mu = 0$ component of a 4-vector is often called 'time component', and the $\mu = 1, 2, 3$ components 'space components.'

Define the inner product (or 'dot' product) $A \cdot B$ of two 4-vectors $A^\mu = (A^0, \vec{A})$ and $B^\mu = (B^0, \vec{B})$ by

$$A \cdot B \stackrel{\text{def}}{\equiv} A^0 B^0 - \vec{A} \cdot \vec{B} = A^0 B^0 - A^1 B^1 - A^2 B^2 - A^3 B^3 . \tag{1.12}$$

Then, $P^2 \equiv P \cdot P$ is nothing but m^2:

$$P^2 = P^{0\,2} - \vec{P}^2 = E^2 - \vec{P}^2 = m^2 \tag{1.13}$$

which is invariant under Lorentz boost. This inner product $P \cdot P$ is similar to the norm $\vec{x}^2$ of an ordinary 3-dimensional vector $\vec{x}$, which is invariant under rotation, except for the minus signs for the space components in the definition of the inner product. In order to handle these minus signs conveniently, we define 'subscripted' components of a 4-vector by

$$A_0 = A^0, \quad A_i = -A^i \ (i = 1, 2, 3) . \tag{1.14}$$

Then the inner product (1.12) can be written as

$$A \cdot B = A_0 B^0 + A_1 B^1 + A_2 B^2 + A_3 B^3 \stackrel{\text{def}}{\equiv} A_\mu B^\mu = A^\mu B_\mu \,, \tag{1.15}$$

where we have used the convention that when a pair of the same index appears in the *same term*, then summation over all possible values of the index ($\mu = 0, 1, 2, 3$ in this case) is implied. In general, we will use Roman letters for space indices (take values 1,2,3) and greek letters for space-time (Lorentz) indices (take values 0,1,2,3). Thus,

$$x^i y^i = \sum_{i=1}^{3} x^i y^i \; (= \vec{x} \cdot \vec{y}), \qquad (A^\mu + B^\mu) C_\mu = \sum_{\mu=0}^{3} (A^\mu + B^\mu) C_\mu \,, \tag{1.16}$$

but no sum over μ or ν in

$$A_\mu B^\nu + C_\mu D^\nu \quad (\mu,\, \nu \text{ not in the same term}). \tag{1.17}$$

When a pair of Lorentz indices is summed over, usually one index is subscript and the other is superscript. Such indices are said to be 'contracted'. Also, it is important that there is only one pair of a given index per term. We do not consider implied summations such as $A^\mu B^\mu C_\mu$ to be well-defined. [$(A^\mu + B^\mu) C_\mu$ is well-defined since it is equal to $A^\mu C_\mu + B^\mu C_\mu$.]

Now, define the metric tensor $g_{\mu\nu}$ by

$$g_{00} = -g_{11} = -g_{22} = -g_{33} = 1, \quad g_{\mu\nu} = 0 \; (\mu \neq \nu) \tag{1.18}$$

which is symmetric:

$$g_{\mu\nu} = g_{\nu\mu} \,. \tag{1.19}$$

The corresponding matrix G is defined as

$$\{g_{\mu\nu}\} \stackrel{\text{def}}{\equiv} \begin{array}{c c} & \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \; = \nu \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ \| \\ \mu \end{array} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \end{array} \stackrel{\text{def}}{\equiv} G \tag{1.20}$$

When we form a matrix out of a quantity with two indices, by definition we take the first index to increase downward, and the second to increase to the right.

As defined in (1.14) for a 4-vector, switching an index between superscript and subscript results in a sign change when the index is 1,2, or 3, while the sign is unchanged when the index is zero. We adopt the same rule for the indices of $g_{\mu\nu}$. In fact, from now on, we enforce the same rule for all space-time indices (unless otherwise stated, such as for the Kronecker delta below). Then we have

$$g_{\mu\nu} = g^{\mu\nu}, \quad g_\mu{}^\nu = g^\mu{}_\nu = \delta_{\mu\nu} \tag{1.21}$$

where $\delta_{\mu\nu}$ is the Kronecker's delta ($\delta_{\mu\nu} = 1$ if $\mu = \nu$, 0 otherwise) which we define to have only subscripts. Then, $g_{\mu\nu}$ can be used together with contraction to 'lower' or 'raise' indices:

$$A_\nu = g_{\mu\nu} A^\mu, \quad A^\nu = g^{\mu\nu} A_\mu \tag{1.22}$$

which are equivalent to the rule (1.14).

The inner product of 4-vectors A and B (1.12) can also be written in matrix form as

$$A \cdot B = A^\mu g_{\mu\nu} B^\nu = \overset{(A^0\ A^1\ A^2\ A^3)}{} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} B^0 \\ B^1 \\ B^2 \\ B^3 \end{pmatrix} = A^T G B\,. \tag{1.23}$$

When we use 4-vectors in matrix form, they are understood to be column vectors with *superscripts*, while their transpose are row vectors:

$$A \overset{\text{def}}{\equiv} \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix} = \begin{pmatrix} A^0 \\ A_x \\ A_y \\ A_z \end{pmatrix}, \quad A^T = (A^0, A^1, A^2, A^3) \qquad \text{(in matrix form)}. \tag{1.24}$$

1.3 Lorentz group

The Lorentz boost (1.4) can be written in matrix form as

$$P' = \Lambda P \tag{1.25}$$

with

$$P' = \begin{pmatrix} E' \\ P'_x \\ P'_y \\ P'_z \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \gamma & \eta & 0 & 0 \\ \eta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} E \\ P_x \\ P_y \\ P_z \end{pmatrix}. \tag{1.26}$$

In terms of components, this can be written as

$$\boxed{P'^\mu = \Lambda^\mu{}_\nu P^\nu}, \tag{1.27}$$

where we have defined the components of the matrix Λ by taking the first index to be superscript and the second to be subscript (still the first index increases downward and the second index increases to the right):

$$\Lambda \stackrel{\text{def}}{\equiv} \{\Lambda^\mu{}_\nu\} = \begin{array}{c c} & \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \quad = \nu \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ \parallel \\ \mu \end{array} & \begin{array}{c} \begin{pmatrix} \gamma & \eta & 0 & 0 \\ \eta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ \\ \\ \end{array} \end{array} \tag{1.28}$$

For example, $\Lambda^0{}_1 = \eta$ and thus $\Lambda^{01} = -\eta$, etc. The superscript and subscript in (1.27) were chosen such that the index ν is contracted and that the index μ on both sides of the equality has consistent position, namely, both are superscript.

We have seen that $P^2 = E^2 - \vec{P}^2$ is invariant under the Lorentz boost given by (1.4) or (1.9). We will now find the necessary and sufficient condition for a 4×4 matrix Λ to leave the inner product of two 4-vectors invariant. Suppose A^μ and B^μ transform by the same matrix Λ:

$$A'^\mu = \Lambda^\mu{}_\alpha A^\alpha, \quad B'^\nu = \Lambda^\nu{}_\beta B^\beta\,. \tag{1.29}$$

Then the inner products $A' \cdot B'$ and $A \cdot B$ can be written using (1.22) as

$$\begin{aligned} A' \cdot B' &= \underbrace{A'_\nu}_{g_{\mu\nu} \underbrace{A'^\mu}_{\Lambda^\mu{}_\alpha A^\alpha}} \underbrace{B'^\nu}_{\Lambda^\nu{}_\beta B^\beta} = (g_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta) A^\alpha B^\beta \\ A \cdot B &= \underbrace{A_\beta}_{g_{\alpha\beta} A^\alpha} B^\beta = g_{\alpha\beta} A^\alpha B^\beta\,. \end{aligned} \tag{1.30}$$

In order for $A' \cdot B' = A \cdot B$ to hold for any A and B, the coefficients of $A^\alpha B^\beta$ should be the same term by term (To see this, set $A^\nu = 1$ for $\nu = \alpha$ and 0 for all else, and $B^\nu = 1$ for $\nu = \beta$ and 0 for all else.):

$$\boxed{g_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = g_{\alpha\beta}}\,. \tag{1.31}$$

On the other hand, if Λ satisfies this condition, the same derivation above can be traced backward to show that the inner product $A \cdot B$ defined by (1.12) is invariant. Thus, (1.31) is the necessary and sufficient condition.

What does the condition (1.31) tell us about the nature of the matrix Λ? Using (1.22), we have $g_{\mu\nu}\Lambda^\mu{}_\alpha = \Lambda_{\nu\alpha}$, then the condition becomes

$$\Lambda_{\nu\alpha}\Lambda^\nu{}_\beta = g_{\alpha\beta} \quad \xrightarrow{\text{raise } \alpha \text{ on both sides}} \quad \Lambda_\nu{}^\alpha\Lambda^\nu{}_\beta = g^\alpha{}_\beta \; (= \delta_{\alpha\beta})\,. \tag{1.32}$$

Comparing this with the definition of the inverse transformation Λ^{-1}:

$$(\Lambda^{-1})^\alpha{}_\nu \Lambda^\nu{}_\beta = \delta_{\alpha\beta}\,, \tag{1.33}$$

we see that the inverse matrix of Λ is obtained by

$$\boxed{(\Lambda^{-1})^\alpha{}_\nu = \Lambda_\nu{}^\alpha}\,, \tag{1.34}$$

which means that one simply has to change the sign of the components for which only one of the indices is zero (namely, $\Lambda^0{}_i$ and $\Lambda^i{}_0$) and then transpose it:

$$\Lambda = \begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{pmatrix}, \quad \longrightarrow \Lambda^{-1} = \begin{pmatrix} \Lambda^0{}_0 & -\Lambda^1{}_0 & -\Lambda^2{}_0 & -\Lambda^3{}_0 \\ -\Lambda^0{}_1 & \Lambda^1{}_1 & \Lambda^2{}_1 & \Lambda^3{}_1 \\ -\Lambda^0{}_2 & \Lambda^1{}_2 & \Lambda^2{}_2 & \Lambda^3{}_2 \\ -\Lambda^0{}_3 & \Lambda^1{}_3 & \Lambda^2{}_3 & \Lambda^3{}_3 \end{pmatrix}. \tag{1.35}$$

Thus, the set of matrices that keep the inner product of 4-vectors invariant is made of matrices that become their own inverse when the signs of components with one time index are flipped and then transposed. As we will see below, such set of matrices forms a group, called the *Lorentz group*, and any such transformation [namely, one that keeps the 4-vector inner product invariant, or equivalently that satisfies the condition (1.31)] is defined as a *Lorentz transformation*.

To show that such set of matrices forms a group, it is convenient to write the condition (1.31) in matrix form. Noting that when written in terms of components, we can change the ordering in any way we want, the condition can be written as

$$\Lambda^\mu{}_\alpha g_{\mu\nu}\Lambda^\nu{}_\beta = g_{\alpha\beta}, \quad \text{or} \quad \Lambda^T G \Lambda = G\,. \tag{1.36}$$

A set forms a group when for any two elements of the set x_1 and x_2, a 'product' x_1x_2 can be defined such that

1. (Closure) The product x_1x_2 also belongs to the set.

2. (Associativity) For any elements x_1, x_2 and x_3, $(x_1x_2)x_3 = x_1(x_2x_3)$.

3. (Identity) There exists an element I in the set that satisfies $Ix = xI = x$ for any element x.

4. (Inverse) For any element x, there exists an element x^{-1} in the set that satisfies $x^{-1}x = xx^{-1} = I$.

In our case at hand, the set is all 4×4 matrices that satisfy $\Lambda^T G \Lambda = G$, and we take the ordinary matrix multiplication as the 'product' which defines the group. The proof is straightforward:

1. Suppose Λ_1 and Λ_2 belong to the set (i.e. $\Lambda_1^T G \Lambda_1 = G$ and $\Lambda_2^T G \Lambda_2 = G$). Then,
$$(\Lambda_1\Lambda_2)^T G (\Lambda_1\Lambda_2) = \Lambda_2^T \underbrace{\Lambda_1^T G \Lambda_1}_{G} \Lambda_2 = \Lambda_2^T G \Lambda_2 = G. \tag{1.37}$$
Thus, the product $\Lambda_1\Lambda_2$ also belongs to the set.

2. The matrix multiplication is of course associative: $(\Lambda_1\Lambda_2)\Lambda_3 = \Lambda_1(\Lambda_2\Lambda_3)$.

3. The identity matrix I ($I^\mu{}_\nu = \delta_{\mu\nu}$) belongs to the set ($I^T G I = G$), and satisfies $I\Lambda = \Lambda I = \Lambda$ for any element.

4. We have already seen that if a 4×4 matrix Λ satisfies $\Lambda^T G \Lambda = G$, then its inverse exists as given by (1.34). It is instructive, however, to prove it more formally. Taking the determinant of $\Lambda^T G \Lambda = G$,
$$\underbrace{\det \Lambda^T}_{\det \Lambda} \underbrace{\det G}_{-1} \det \Lambda = \underbrace{\det G}_{-1} \quad \to \quad (\det \Lambda)^2 = 1\,, \tag{1.38}$$
where we have used the property of determinant
$$\det(MN) = \det M \det N \tag{1.39}$$
with M and N being square matrices of same rank. Thus, $\det \Lambda \neq 0$ and therefore its inverse Λ^{-1} exists. Also, it belongs to the set: multiplying $\Lambda^T G \Lambda = G$ by $(\Lambda^{-1})^T$ from the left and by Λ^{-1} from the right,
$$\underbrace{(\Lambda^{-1})^T \Lambda^T}_{(\underbrace{\Lambda\Lambda^{-1}}_{I})^T} G \underbrace{\Lambda\Lambda^{-1}}_{I} = (\Lambda^{-1})^T G \Lambda^{-1} \ \to\ (\Lambda^{-1})^T G \Lambda^{-1} = G\,. \tag{1.40}$$
This completes the proof that Λ's that satisfy (1.36) form a group.

Since the inverse of a Lorentz transformation is also a Lorentz transformation, it should satisfy the condition (1.31)
$$g_{\alpha\beta} = g_{\mu\nu} (\Lambda^{-1})^\mu{}_\alpha (\Lambda^{-1})^\nu{}_\beta = g_{\mu\nu} \Lambda_\alpha{}^\mu \Lambda_\beta{}^\nu \quad \to \quad \boxed{g_{\mu\nu} \Lambda_\alpha{}^\mu \Lambda_\beta{}^\nu = g_{\alpha\beta}}\,, \tag{1.41}$$

where we have used the inversion rule (1.34). The formulas (1.31), (1.41), and their variations are then summarized as follows: on the left hand side of the form $g\,\Lambda\,\Lambda = g$, an index of g (call it μ) is contracted with an index of a Λ and the other index of g (call it ν) with an index of the other Λ. As long as μ and ν are both first or both second indices on the Λ's, and as long as the rest of the indices are the same (including superscript/subscript) on both sides of the equality, any possible way of indexing gives a correct formula. Similarly, on the left hand side of the form $\Lambda\,\Lambda = g$, an index of a Λ and an index of the other Λ are contracted. As long as the contracted indices are both first or both second indices on Λ's, and as long as the rest of the indices are the same on both sides of the equality, any possible way of indexing gives a correct formula.

A natural question at this point is whether the Lorentz group defined in this way is any larger than the set of Lorentz boosts defined by (1.9). The answer is yes. Clearly, any rotation in the 3-dimensional space keeps $\vec{A}\cdot\vec{B}$ invariant while it does not change the time components A^0 and B^0. Thus, it keeps the 4-vector inner product $A\cdot B = A^0B^0 - \vec{A}\cdot\vec{B}$ invariant, and as a result it belongs to the Lorentz group by definition. On the other hand, the only way the boost (1.9) does not change the time component is to set $\beta = 0$ in which case the transformation is the identity transformation. Thus, any finite rotation in the 3-dimensional space is not a boost.

Furthermore, the time reversal T and the space inversion P defined by

$$T \stackrel{\text{def}}{\equiv} \{T^\mu{}_\nu\} \stackrel{\text{def}}{\equiv} \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}, \quad P \stackrel{\text{def}}{\equiv} \{P^\mu{}_\nu\} \stackrel{\text{def}}{\equiv} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{1.42}$$

satisfy

$$T^T G T = G, \quad P^T G P = G\,, \tag{1.43}$$

and thus belong to the Lorentz group. Even though the matrix P has the same numerical form as G, it should be noted that P is a Lorentz transformation but G is not (it is a metric). The difference is also reflected in the fact that the matrix P is defined by the first index being superscript and the second subscript (because it is a Lorentz transformation), while the matrix G is defined by both indices being subscript (or both superscript).

As we will see later, boosts and rotations can be formed by consecutive infinitesimal transformations starting from identity I (they are 'continuously connected' to I), while T and P cannot (they are 'disconnected' from I, or said to be 'discrete' transformations). Any product of boosts, rotation, T, and P belongs to the Lorentz group, and it turns out that they saturate the Lorentz group. Thus, we write symbolically

$$\text{Lorentz group} \;=\; \text{boost} \;+\; \text{rotation} \;+T+P\,. \tag{1.44}$$

Later, we will see that any Lorentz transformation continuously connected to I is a boost, a rotation, or a combination thereof.

If the origins of the inertial frames K and K' touch at $t = t' = 0$ and $\vec{x} = \vec{x}' = 0$, the coordinate $x^\mu = (t, \vec{x})$ of any event transforms in the same way as P^μ:

$$x^\mu = \Lambda^\mu{}_\nu x^\nu \, . \tag{1.45}$$

This can be extended to include space-time translation between the two frames:

$$x^\mu = \Lambda^\mu{}_\nu x^\nu \ + a^\mu \, , \tag{1.46}$$

where a^μ is a constant 4-vector. The transformation of energy-momentum is not affected by the space-time translation, and is still given by $P' = \Lambda P$. Such transformations that include space-time translation also form a group and called the 'inhomogeneous Lorentz group' or the 'Poincaré group'. The group formed by the transformations with $a^\mu = 0$ is sometimes called the homogeneous Lorentz group. Unless otherwise stated, we will deal with the homogeneous Lorentz group; namely without space-time translation.

1.4 Classification of Lorentz transformations

Up to this point, we have not specified that Lorentz transformations are real (namely, all the elements are real). In fact, Lorentz transformations as defined by (1.31) in general can be complex and the complex Lorentz transformations plays an important role in a formal proof of an important symmetry theorem called CPT theorem which states that the laws of physics are invariant under the combination of particle-antiparticle exchange (C), mirror inversion (P), and time reversal (T) under certain natural assumptions. In this book, however, we will assume that Lorentz transformations are real.

As seen in (1.38), all Lorentz transformation satisfy $(\det \Lambda)^2 = 1$, or equivalently, $\det \Lambda = +1$ or -1. We define 'proper' and 'improper' Lorentz transformations as

$$\begin{cases} \det \Lambda & = +1 \ : \quad \text{proper} \\ \det \Lambda & = -1 \ : \quad \text{improper} \end{cases} \, . \tag{1.47}$$

Since $\det(\Lambda_1 \Lambda_2) = \det \Lambda_1 \det \Lambda_2$, the product of two proper transformations or two improper transformations is proper, while the product of a proper transformation and a improper transformation is improper.

Next, look at the $(\alpha, \beta) = (0, 0)$ component of the defining condition $g_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = g_{\alpha\beta}$:

$$g_{\mu\nu} \Lambda^\mu{}_0 \Lambda^\nu{}_0 = g_{00} = 1 \, , \quad \to \quad (\Lambda^0{}_0)^2 - \sum_{i=1}^{3} (\Lambda^i{}_0)^2 = 1 \tag{1.48}$$

	$\Lambda^0{}_0 \geq 1$ orthochronous	$\Lambda^0{}_0 \leq -1$ non-orthochronous
$\det \Lambda = +1$ proper	$\Lambda^{(\mathrm{po})}$	$TP\Lambda^{(\mathrm{po})}$
$\det \Lambda = -1$ improper	$P\Lambda^{(\mathrm{po})}$	$T\Lambda^{(\mathrm{po})}$

Table 1.1: Classification of the Lorentz group. $\Lambda^{(\mathrm{po})}$ is any proper and orthochronous Lorentz transformation.

or

$$(\Lambda^0{}_0)^2 = 1 + \sum_{i=1}^{3} (\Lambda^i{}_0)^2 \geq 1\,, \tag{1.49}$$

which means $\Lambda^0{}_0 \geq 1$ or $\Lambda^0{}_0 \leq -1$, and this defines the 'orthochronous' and 'non-orthochronous' Lorentz transformations:

$$\begin{cases} \Lambda^0{}_0 \ \geq 1 \ : & \text{orthochronous} \\ \Lambda^0{}_0 \ \leq -1 \ : & \text{non-orthochronous} \end{cases} . \tag{1.50}$$

It is easy to show that the product of two orthochronous transformations or two non-orthochronous transformations is orthochronous, and the product of an orthochronous transformation and a non-orthochronous transformation is non-orthochronous.

From the definitions (1.42) and $I^\mu{}_\nu = \delta_{\mu\nu}$, we have

$$\begin{gathered} \det I = \det(TP) = +1, \ \ \det T = \det P = -1, \\ I^0{}_0 = P^0{}_0 = +1, \qquad T^0{}_0 = (TP)^0{}_0 = -1 \end{gathered} \tag{1.51}$$

Thus, the identity I is proper and orthochronous, P is improper and orthochronous, T is improper and non-orthochronous, and TP is proper and non-orthochronous. Accordingly, we can multiply any proper and orthochronous transformations by each of these to form four sets of transformations of given properness and orthochronousness as shown in Table 1.1. Any Lorentz transformation is proper or improper (i.e. $\det \Lambda = \pm 1$) and orthochronous or non-orthochronous (i.e. $|\Lambda^0{}_0|^2 \geq 1$). Since any transformation that is not proper and orthochronous can be made proper and orthochronous by multiplying T, P or TP, the four forms of transformations in Table 1.1 saturate the Lorentz group. For example, if Λ is improper and orthochronous, then $P\Lambda \overset{\text{def}}{\equiv} \Lambda^{(\mathrm{po})}$ is proper and orthochronous, and Λ can be written as $\Lambda = PP\Lambda = P\Lambda^{(\mathrm{po})}$.

It is straightforward to show that the set of proper transformations and the set of orthochronous transformations separately form a group, and that proper and orthochronous transformations by themselves form a group. Also, the set of proper and orthochronous transformations and the set of improper and non-orthochronous transformations together form a group.

Exercise 1.1 *Classification of Lorentz transformations.*
(*a*) *Suppose* $\Lambda = AB$ *where* $\Lambda, A,$ *and* B *are Lorentz transformations. Prove that* Λ *is orthochronous if* A *and* B *are both orthochronous or both non-orthochronous, and that* Λ *is non-orthochronous if one of* A *and* B *is orthochronous and the other is non-orthochronous.*
[*hint: Note that we can write* ${\Lambda^0}_0 = {A^0}_0{B^0}_0 + \vec{a}\cdot\vec{b}$ *with* $\vec{a} \equiv ({A^0}_1, {A^0}_2, {A^0}_3)$ *and* $\vec{b} \equiv ({B^1}_0, {B^2}_0, {B^3}_0)$. *Then use* $|\vec{a}\cdot\vec{b}| \leq |\vec{a}||\vec{b}|$. *Also, one can derive* $\vec{a}^2 = {{A^0}_0}^2 - 1$ *and* $\vec{b}^2 = {{B^0}_0}^2 - 1$.]
(*b*) *Show that the following sets of Lorentz transformations each form a group:*

1. *proper transformations*
2. *orthochronous transformations*
3. *proper and orthochronous transformations*
4. *proper and orthochronous transformations plus improper and non-orthochronous transformations*

As mentioned earlier (and as will be shown later) boosts and rotations are continuously connected to the identity. Are they then proper and orthochronous? To show that this is the case, it suffices to prove that an infinitesimal transformation can change $\det\Lambda$ and ${\Lambda^0}_0$ only infinitesimally, since then multiplying an infinitesimal transformation cannot jump across the gap between $\det\Lambda = +1$ and $\det\Lambda = -1$ or the gap between ${\Lambda^0}_0 \geq 1$ and ${\Lambda^0}_0 \leq -1$.

An infinitesimal transformation is a transformation that is very close to the identity I and any such transformation λ can be written as

$$\lambda = I + dH \tag{1.52}$$

where d is a small number and H is a 4×4 matrix of order unity meaning the maximum of the absolute values of its elements is about 1. To be specific, we could define it such that $\max_{\alpha,\beta}|{H^\alpha}_\beta| = 1$ and $d \geq 0$, which uniquely defines the decomposition above. We want to show that for any Lorentz transformation Λ, multiplying $I + dH$ changes the determinant or the $(0,0)$ component only infinitesimally; namely, the differences vanish as we take d to zero.

The determinant of a $n \times n$ matrix A is defined by

$$\det A \stackrel{\text{def}}{\equiv} \sum_{\text{permutations}} s_{i_1,i_2,\ldots i_n} A_{i_1 1} A_{i_2 2} \ldots A_{i_n n} \tag{1.53}$$

where the sum is taken over $(i_1, i_2, \ldots i_n)$ which is any permutation of $(1, 2, \ldots n)$, and $s_{i_1,i_2,\ldots i_n}$ is $1(-1)$ if $(i_1, i_2, \ldots i_n)$ is an even(odd) permutation. When applied to 4×4 Lorentz transformations, this can be written as

$$\det A \stackrel{\text{def}}{\equiv} \epsilon_{\alpha\beta\gamma\delta} A^\alpha{}_0 A^\beta{}_1 A^\gamma{}_2 A^\delta{}_3 \,, \tag{1.54}$$

where the implicit sum is over $\alpha, \beta, \gamma, \delta = 0, 1, 2, 3$ and $\epsilon_{\alpha\beta\gamma\delta}$ is the totally anti-symmetric 4-th rank tensor defined by

$$\epsilon_{\alpha\beta\gamma\delta} \stackrel{\text{def}}{\equiv} \begin{cases} = \begin{Bmatrix} +1 \\ -1 \end{Bmatrix} & \text{if } (\alpha\beta\gamma\delta) \text{ is an } \begin{Bmatrix} \text{even} \\ \text{odd} \end{Bmatrix} \text{ permutation of } (0,1,2,3) \\ = \;\; 0 & \text{if any of } \alpha\beta\gamma\delta \text{ are equal} \end{cases} \tag{1.55}$$

The standard superscript/subscript rule applies to the indices of $\epsilon_{\alpha\beta\gamma\delta}$; namely, $\epsilon_{0123} = -\epsilon^{0123} = 1$, etc. Then, it is easy to show that

$$\det(I + dH) = 1 + d\,\mathrm{Tr}H + (\text{higher orders in } d)\,, \tag{1.56}$$

where the 'trace' of a matrix A is defined as the sum of the diagonal elements:

$$\mathrm{Tr}A \stackrel{\text{def}}{\equiv} \sum_{\alpha=0}^{3} A^\alpha{}_\alpha \,. \tag{1.57}$$

Exercise 1.2 *Determinant and trace.*
Determinant of a $n \times n$ matrix is defined by

$$\det A \stackrel{\text{def}}{\equiv} s_{i_1 i_2,\ldots i_n} A_{i_1 1} A_{i_2 2} \ldots A_{i_n n}$$

where sum over $i_1, i_2 \ldots i_n$ is implied (each taking values 1 through n) and $s(i_1, i_2 \ldots i_n)$ is the totally asymmetric n-th rank tensor:

$$s_{i_1 i_2 \ldots i_n} \equiv \begin{cases} +1(-1) & \textit{if } (i_1, i_2 \ldots i_n) \textit{ is an even (odd) permutation of } (1, 2, \ldots n). \\ 0 & \textit{if any of } i_1, i_2 \ldots i_n \textit{ are equal.} \end{cases}$$

Show that to first order of a small number d, the determinant of a matrix that is infinitesimally close to the identity matrix I is given by

$$\det(I + dH) = 1 + d\mathrm{Tr}H + (\textit{higher orders in } d)\,,$$

where H is a certain matrix whose size is of order 1, and the trace (Tr) *of a matrix is defined by*

$$\mathrm{Tr}H \equiv \sum_{i=1}^{n} H_{ii} \,.$$

Since all diagonal elements of H are of order unity or smaller, (1.56) tells us that $\det\lambda \to 1$ as we take $d \to 0$. In fact, the infinitesimal transformation λ is a Lorentz transformation, so we know that $\det\lambda = \pm 1$. Thus, we see that the determinant of an infinitesimal transformation is strictly $+1$. It then follows from $\det(\lambda\Lambda) = \det\lambda \det\Lambda$ that multiplying an infinitesimal transformation λ to any transformation Λ does not change the determinant of the transformation.

The $(0,0)$ component of $\lambda\Lambda$ is

$$(\lambda\Lambda)^0{}_0 = [(I + dH)\Lambda]^0{}_0 = [\Lambda + dH\Lambda]^0{}_0 = \Lambda^0{}_0 + d\,(H\Lambda)^0{}_0\,. \tag{1.58}$$

Since $(H\Lambda)^0{}_0$ is a finite number for a finite Λ, the change in the $(0,0)$ component tends to zero as we take $d \to 0$. Thus, no matter how many infinitesimal transformations are multiplied to Λ, the $(0,0)$ component cannot jump across the gap between $+1$ and -1.

Thus, continuously connected Lorentz transformations have the same 'properness' and 'orthochronousness. Therefore, boosts and rotations, which are continuously connected to the identity, are proper and orthochronous.

Do Lorentz boosts form a group?
A natural question is whether Lorentz boosts form a group by themselves. The answer is no, and this is because two consecutive boosts in different directions turn out to be a boost *plus* a rotation as we will see when we study the generators of the Lorentz group. Thus, boosts and rotations have to be combined to form a group. On the other hand, rotations form a group by themselves.

1.5 Tensors

Suppose A^μ and B^μ are 4-vectors. Each is a set of 4 numbers that transform under a Lorentz transformation Λ as

$$A'^\mu = \Lambda^\mu{}_\alpha A^\alpha, \quad B'^\nu = \Lambda^\nu{}_\beta B^\beta\,. \tag{1.59}$$

Then, the set of 16 numbers $A^\mu B^\nu$ $(\mu,\nu = 0,1,2,3)$ transforms as

$$A'^\mu B'^\nu = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta A^\alpha B^\beta\,. \tag{1.60}$$

Anything that has 2 Lorentz indices, which is a set of 16 numbers, and transforms as

$$(\quad)'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta (\quad)^{\alpha\beta} \tag{1.61}$$

is called a second rank tensor (or simply a 'tensor'). It may be real, complex, or even a set of operators. Similarly, a quantity that has 3 indices and transforms as

$$(\quad)'^{\mu\nu\sigma} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \Lambda^\sigma{}_\gamma (\quad)^{\alpha\beta\gamma} \tag{1.62}$$

is called a third-rank tensor, and so on. A 4-vector (or simply a 'vector') is a first-rank tensor. A Lorentz-invariant quantity, sometimes called a 'scalar', has no Lorentz index, and thus it is a zero-th rank tensor:

$$(\quad)' = (\quad) \qquad (\text{scalar})\,. \tag{1.63}$$

Contracted indices do not count in deciding the rank of a tensor. For example,

$$A^\mu B_\mu \ :\ (\text{scalar}), \quad A_\mu T^{\mu\nu} \ :\ (\text{vector}), \quad F^{\mu\nu} G_{\mu\sigma} \ :\ (\text{tensor}), \quad \text{etc.} \tag{1.64}$$

The metrix $g_{\mu\nu}$ has two Lorentz indices and thus can be considered a second-rank tensor (thus, the metric *tensor*), then it should transform as

$$g'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta g^{\alpha\beta} = g^{\mu\nu} \tag{1.65}$$

where the second equality is due to (1.31). Namely, the metric tensor is invariant under Lorentz transformations.

In order for some equation to be *Lorentz-invariant*, the Lorentz indices have to be the same on both sides of the equality, including the superscript/subscript distinction. By 'Lorentz-invariant', we mean that if an equation holds in one frame, then it holds in any other frame after all the quantities that appear in the equation are evaluated in the new frame. In the literature, such equations are sometimes called *Lorentz covariant*: both sides of the equality change values but the form stays the same. For example, if an equation $A^{\mu\nu} = B^{\mu\nu}$ (which is actually a set of 16 equations) holds in a frame, then it also holds in any other frame:

$$A'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta A^{\alpha\beta} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta B^{\alpha\beta} = B'^{\mu\nu}\,. \tag{1.66}$$

Thus, equations such as

$$\begin{aligned} m^2 &= P^\mu P_\mu\,, \\ P^\mu &= A^\mu + B^\mu\,, \\ F^{\mu\nu} &= A^\mu B^\nu \end{aligned} \tag{1.67}$$

are all Lorentz-invariant, assuming of course that the quantities transform in the well-defined ways as described above.

1.6 Fields (classical)

A field is a quantity that is a function of space-time point $x^\mu = (t, \vec{x})$ (or 'event'). A scalar quantity that is a function of space time is called a scalar field, a vector quantity that is a function of space time is called a vector field, etc. The rank of a field and the Lorentz transformation properties (scalar, vector, tensor, etc.) are defined in the same way as before, provided that the quantities are evaluated at the same event point before and after a Lorentz transformation; namely,

$$\boxed{\begin{array}{lll} \text{Scalar field}: & \phi'(x') & = \phi(x) \\ \text{Vector field}: & A'^{\mu}(x') & = \Lambda^{\mu}{}_{\alpha} A^{\alpha} \\ \text{Tensor field}: & T'^{\mu\nu}(x') & = \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta} T^{\alpha\beta}(x) \end{array}} \tag{1.68}$$

where x' and x are related by

$$x'^{\mu} = \Lambda^{\mu}{}_{\alpha} x^{\alpha}\,. \tag{1.69}$$

For example, a vector field associates a set of 4 numbers $A^{\mu}(x)$ to an event point x, say when an ant sneezes. In another frame, there are a set of 4 numbers $A'^{\mu}(x')$ associated with the *same event* x', namely, when the ant sneezes in *that* frame, and they are related to the 4 numbers $A^{\mu}(x)$ in the original frame by the matrix Λ. The functional shape of a primed field is in general different from that of the corresponding unprimed field. Namely, if one plots $\phi(x)$ as a function of x and $\phi'(x')$ as a function of x', they will look different.

When a quantity is a function of x, we naturally encounter space-time derivatives of such quantity. Then a question arises as to how they transform under a Lorentz transformation. Take a scalar field $f(x)$, and form a set of 4 numbers (fields) by taking space-time derivatives:

$$\begin{aligned} \frac{\partial f}{\partial x^{\mu}}(x) &= \left(\frac{\partial f}{\partial x^0}(x), \frac{\partial f}{\partial x^1}(x), \frac{\partial f}{\partial x^2}(x), \frac{\partial f}{\partial x^3}(x)\right) \\ &= \left(\frac{\partial f}{\partial t}(x), \frac{\partial f}{\partial x}(x), \frac{\partial f}{\partial y}(x), \frac{\partial f}{\partial z}(x)\right)\,. \end{aligned} \tag{1.70}$$

Then pick two space-time points x_1 and x_2 which are close in space and in time. The argument below is based on the observation that the difference between the values of the scalar field at the two event points is the same in any frame. Since $f(x)$ is a scalar field, the values at a given event is the same before and after a Lorentz transformation:

$$f'(x'_1) = f(x_1), \qquad f'(x'_2) = f(x_2)\,, \tag{1.71}$$

or

$$f'(x'_1) - f'(x'_2) = f(x_1) - f(x_2)\,. \tag{1.72}$$

Since x_1 and x_2 are close, this can be written as

$$dx'^{\mu}\frac{\partial f'}{\partial x'^{\mu}}(x'_1) = dx^{\mu}\frac{\partial f}{\partial x^{\mu}}(x_1)\,, \tag{1.73}$$

where summation over μ is implied, and

$$dx'^{\mu} \stackrel{\text{def}}{\equiv} x'^{\mu}_1 - x'^{\mu}_2\,, \quad dx^{\mu} \stackrel{\text{def}}{\equiv} x^{\mu}_1 - x^{\mu}_2\,. \tag{1.74}$$

which tells us that the quantity $dx^\mu(\partial f/\partial x^\mu)$ is Lorentz-invariant. Since $dx^\mu = x_1^\mu - x_2^\mu$ is a superscripted 4-vector, it follows that $\partial f/\partial x^\mu$ should transform as a subscripted 4-vector (which transforms as $A_\mu = \Lambda_\mu{}^\alpha A_\alpha$):

$$\frac{\partial f'}{\partial x'^\mu}(x') = \Lambda_\mu{}^\alpha \frac{\partial f}{\partial x^\alpha}(x)\,. \tag{1.75}$$

In fact, together with $dx'^\mu = \Lambda^\mu{}_\beta dx^\beta$, we have

$$\begin{aligned} dx'^\mu \frac{\partial f'}{\partial x'^\mu}(x') &= \Big(\Lambda^\mu{}_\beta dx^\beta\Big)\left(\Lambda_\mu{}^\alpha \frac{\partial f}{\partial x^\alpha}(x)\right) \\ &= \underbrace{\Lambda_\mu{}^\alpha \Lambda^\mu{}_\beta}_{g^\alpha{}_\beta \quad \text{by (1.32)}} dx^\beta \frac{\partial f}{\partial x^\alpha}(x) \\ &= dx^\alpha \frac{\partial f}{\partial x^\alpha}(x)\,, \end{aligned} \tag{1.76}$$

showing that it is indeed Lorentz-invariant.

Thus, the index μ in the differential operator $\partial/\partial x^\mu$ acts as a subscript even though it is a superscript on x. To make this point clear, $\partial/\partial x^\mu$ is often written using a subscript as

$$\partial_\mu \stackrel{\text{def}}{\equiv} \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \left(\frac{\partial}{\partial t}, \vec\nabla\right)\,. \tag{1.77}$$

Once ∂_μ is defined, the standard subscript/superscript rule applies; namely, $\partial^\mu = \partial/\partial x_\mu$, etc. Symbolically, the operator ∂^μ then transforms as a superscripted 4-vector:

$$\partial'^\mu = \Lambda^\mu{}_\nu \partial^\nu\,, \tag{1.78}$$

with $\partial'^\mu \equiv \partial/\partial x'_\mu$.

Example: Consider the 4-component charge current density $j^\mu(x) = (\rho(x), \vec j(x))$. We can see that this is indeed a Lorentz 4-vector as follows: Suppose the charge is carried by some medium, such as gas of ions, then pick a space-time point x and let ρ_0 be the charge density in the rest frame of the medium and $\vec\beta$ be the velocity of the medium at that point. Then the charge density ρ in the frame in question is larger than ρ_0 by the factor $\gamma = 1/\sqrt{1-\beta^2}$ due to Lorentz contraction

$$\rho = \rho_0\gamma\,. \tag{1.79}$$

Since $\vec j = \rho\vec\beta$, j^μ can be written as

$$j^\mu = (\rho, \vec j) = (\rho_0\gamma, \rho_0\gamma\vec\beta) = \rho_0(\gamma, \vec\eta) = \rho_0\eta^\mu\,, \tag{1.80}$$

where we have defined the '4-velocity' η^μ by

$$\eta^\mu \stackrel{\text{def}}{\equiv} (\gamma, \gamma\vec{\beta})\,. \tag{1.81}$$

On the other hand, the 4-momentum of a particle with mass m can be written as

$$P^\mu = (m\gamma, m\gamma\vec{\beta}) = m\eta^\mu \tag{1.82}$$

which means that the 4-velocity η^μ is a Lorentz 4-vector, and therefore so is j^μ. When the charge is carried by more than one different media, unique rest frame of the media where ρ^0 is defined does not exist. The total j^μ, however, is the sum of j^μ for each medium. Since j^μ for each medium is a 4-vector, the sum is also a 4-vector.

Then $\partial_\mu j^\mu = 0$ is a Lorentz-invariant equation; namely, if it is true in one frame, then it is true in any frame. Using (1.77), we can write $\partial_\mu j^\mu = 0$ as

$$\begin{aligned}
\partial_\mu j^\mu &= \partial_0 j^0 + \partial_1 j^1 + \partial_2 j^2 + \partial_3 j^3 \\
&= \frac{\partial}{\partial t}\rho + \frac{\partial}{\partial x} j_x + \frac{\partial}{\partial y} j_y + \frac{\partial}{\partial z} j_z \\
&= \frac{\partial}{\partial t}\rho \underset{\substack{\uparrow \\ \text{note the sign!}}}{+} \vec{\nabla}\cdot\vec{j} = 0\,,
\end{aligned} \tag{1.83}$$

which is nothing but the charge conservation equation. Thus, we see that if charge is conserved in one frame it is conserved in any frame.

1.7 Generators of the Lorentz group

In this section, we will focus on the proper and orthochronous Lorentz group. Other elements of the Lorentz group can be obtained by multiplying T, P, and TP to the elements of this group. The goal is to show that any element Λ that is continuously connected to the identity can be written as[2]

$$\Lambda = e^{\xi_i K_i + \theta_i L_i}, \quad (i = 1, 2, 3) \tag{1.84}$$

where ξ_i and θ_i are real numbers and K_i and L_i are 4×4 matrices. Such group whose elements can be parametrized by a set of continuous real numbers (in our case they

[2]In the literature, it is often defined as $\exp i(\xi_i K_i + \theta_i L_i)$, which would make the operators hermitian if the transformation were unitary (e.g. representations of the Lorentz group in the Hilbert space). The Lorentz transformation matrices in space-time are in general not unitary, and for now, we will define without the 'i' so that the expressions become simpler.

are ξ_i and θ_i) is called a *Lie group*. The operators K_i and L_i are called the *generators* of the Lie group.

Any element of the proper and orthochronous Lorentz group is continuously connected to the identity. Actually we have not proven this, but we will at least show that all boosts, rotations and combinations thereof are continuously connected to the identity (and vice versa).

1.7.1 Infinitesimal transformations

Let's start by looking at a Lorentz transformation which is infinitesimally close to the identity:

$$\Lambda^\mu{}_\nu = g^\mu{}_\nu + \omega^\mu{}_\nu \tag{1.85}$$

where $\omega^\mu{}_\nu$ is a set of small (real) numbers. Inserting this to the defining condition (1.31) or equivalently $\Lambda_{\nu\alpha}\Lambda^\nu{}_\beta = g_{\alpha\beta}$ (1.32), we get

$$\begin{aligned} g_{\alpha\beta} &= \Lambda_{\nu\alpha}\Lambda^\nu{}_\beta \\ &= (g_{\nu\alpha} + \omega_{\nu\alpha})(g^\nu{}_\beta + \omega^\nu{}_\beta) \\ &= g_{\nu\alpha}g^\nu{}_\beta + \omega_{\nu\alpha}g^\nu{}_\beta + g_{\nu\alpha}\omega^\nu{}_\beta + \omega_{\nu\alpha}\omega^\nu{}_\beta \\ &= g_{\alpha\beta} + \omega_{\beta\alpha} + \omega_{\alpha\beta} + \omega_{\nu\alpha}\omega^\nu{}_\beta \,. \end{aligned} \tag{1.86}$$

Keeping terms to the first order in ω, we then obtain

$$\omega_{\beta\alpha} = -\omega_{\alpha\beta} \,. \tag{1.87}$$

Namely, $\omega_{\alpha\beta}$ is anti-symmetric (which is true when the indices are both subscript or both superscript; in fact, $\omega^\alpha{}_\beta$ is *not* anti-symmetric under $\alpha \leftrightarrow \beta$), and thus it has 6 independent parameters:

$$\{\omega_{\alpha\beta}\} = \begin{array}{cc} & \beta \longrightarrow \\ \begin{array}{c}\alpha \\ \downarrow\end{array} & \begin{pmatrix} 0 & \boxed{\omega_{01}} & \boxed{\omega_{02}} & \boxed{\omega_{03}} \\ -\omega_{01} & 0 & \boxed{\omega_{12}} & \boxed{\omega_{13}} \\ -\omega_{02} & -\omega_{12} & 0 & \boxed{\omega_{23}} \\ -\omega_{03} & -\omega_{13} & -\omega_{23} & 0 \end{pmatrix} \end{array} \tag{1.88}$$

This can be conveniently parametrized using 6 anti-symmetric matrices as

$$\begin{aligned} \{\omega_{\alpha\beta}\} &= \omega_{01}\{(M^{01})_{\alpha\beta}\} + \omega_{02}\{(M^{02})_{\alpha\beta}\} + \omega_{03}\{(M^{03})_{\alpha\beta}\} \\ &+ \omega_{23}\{(M^{23})_{\alpha\beta}\} + \omega_{13}\{(M^{13})_{\alpha\beta}\} + \omega_{12}\{(M^{12})_{\alpha\beta}\} \\ &= \sum_{\mu<\nu} \omega_{\mu\nu}\{(M^{\mu\nu})_{\alpha\beta}\} \end{aligned} \tag{1.89}$$

with

$$
\begin{aligned}
\{(M^{01})_{\alpha\beta}\} &= \left(\begin{array}{c|ccc} 0 & 1 & 0 & 0 \\ \hline -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right), \quad \{(M^{23})_{\alpha\beta}\} = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{array}\right), \\
\{(M^{02})_{\alpha\beta}\} &= \left(\begin{array}{c|ccc} 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right), \quad \{(M^{13})_{\alpha\beta}\} = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right), \\
\{(M^{03})_{\alpha\beta}\} &= \left(\begin{array}{c|ccc} 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{array}\right), \quad \{(M^{12})_{\alpha\beta}\} = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)
\end{aligned}
\tag{1.90}
$$

Note that for a given pair of μ and ν, $\{(M^{\mu\nu})_{\alpha\beta}\}$ is a 4×4 matrix, while $\omega^{\mu\nu}$ is a real number. The elements $(M^{\mu\nu})_{\alpha\beta}$ can be written in a concise form as follows: first, we note that in the upper right half of each matrix (i.e. for $\alpha < \beta$), the element with $(\alpha, \beta) = (\mu, \nu)$ is 1 and all else are zero, which can be written as $g^{\mu}{}_{\alpha}\, g^{\nu}{}_{\beta}$. For the lower half, all we have to do is to flip α and β and add a minus sign. Combining the two halves, we get

$$(M^{\mu\nu})_{\alpha\beta} = g^{\mu}{}_{\alpha}\, g^{\nu}{}_{\beta} - g^{\mu}{}_{\beta}\, g^{\nu}{}_{\alpha}\,. \tag{1.91}$$

This is defined only for $\mu < \nu$ so far. For $\mu > \nu$, we will use this same expression as the definition; then, $(M^{\mu\nu})_{\alpha\beta}$ is anti-symmetric with respect to $(\mu \leftrightarrow \nu)$:

$$(M^{\mu\nu})_{\alpha\beta} = -(M^{\nu\mu})_{\alpha\beta}\,, \tag{1.92}$$

which also means $(M^{\mu\nu})_{\alpha\beta} = 0$ if $\mu = \nu$. Together with $\omega_{\mu\nu} = -\omega_{\nu\mu}$, (1.89) becomes

$$\omega_{\alpha\beta} = \sum_{\mu<\nu} \omega_{\mu\nu}(M^{\mu\nu})_{\alpha\beta} = \sum_{\mu>\nu} \omega_{\mu\nu}(M^{\mu\nu})_{\alpha\beta} = \frac{1}{2}\omega_{\mu\nu}(M^{\mu\nu})_{\alpha\beta}\,, \tag{1.93}$$

where in the last expression, sum over all values of μ and ν is implied. The infinitesimal transformation (1.85) can then be written as

$$\Lambda^{\alpha}{}_{\beta} = g^{\alpha}{}_{\beta} + \frac{1}{2}\omega_{\mu\nu}(M^{\mu\nu})^{\alpha}{}_{\beta}\,, \tag{1.94}$$

or in matrix form,

$$\Lambda = I + \frac{1}{2}\omega_{\mu\nu} M^{\mu\nu}\,. \tag{1.95}$$

where the first indices of $M^{\mu\nu}$, which is a 4×4 matrix for given μ and ν, is taken to be superscript and the second subscript; namely, in the same way as Lorentz transformation. Namely, when no explicit indexes for elements are given, the 4×4 matrix $M^{\mu\nu}$ is defined as

$$M^{\mu\nu} \stackrel{\text{def}}{\equiv} \{(M^{\mu\nu})^{\alpha}{}_{\beta}\} . \tag{1.96}$$

It is convenient to divide the six matrices to two groups as

$$K_i \stackrel{\text{def}}{\equiv} M^{0i}, \qquad L_i \stackrel{\text{def}}{\equiv} M^{jk} \quad (i,j,k : \text{cyclic}) . \tag{1.97}$$

We always use subscripts for K_i and L_i since only possible values are $i = 1,2,3$, and similarly to $M^{\mu\nu}$, elements of the matrices K_i's and L_i's are defined by taking the first Lorentz index to be superscript and the second subscript:

$$K_i \stackrel{\text{def}}{\equiv} \{(K_i)^{\alpha}{}_{\beta}\} , \quad L_i \stackrel{\text{def}}{\equiv} \{(L_i)^{\alpha}{}_{\beta}\} . \tag{1.98}$$

Later, we will see that K's generate boosts and L's generate rotations. Explicitly, they can be obtained by raising the index α in (1.90) (note also the the minus sign in $L_2 = -M^{13}$):

$$K_1 = \left(\begin{array}{c|ccc} 0 & 1 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right), K_2 = \left(\begin{array}{c|ccc} 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right), K_3 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right) \tag{1.99}$$

$$L_1 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right), L_2 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right), L_3 = \left(\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{1.100}$$

By inspection, we see that the elements of K's and L's can be written as

$$\begin{array}{ll} (K_i)^j{}_k = 0 \,, & (K_i)^0{}_\mu = (K_i)^\mu{}_0 = g^i{}_\mu \,, \\ (L_i)^j{}_k = -\epsilon_{ijk} \,, & (L_i)^0{}_\mu = (L_i)^\mu{}_0 = 0 \,, \end{array} \quad (i,j,k = 1,2,3;\ \mu = 0,1,2,3) \tag{1.101}$$

where ϵ_{ijk} is a totally anti-symmetric quantity defined for $i,j,k = 1,2,3$:

$$\epsilon_{ijk} \stackrel{\text{def}}{\equiv} \begin{cases} = \left\{ \begin{array}{c} +1 \\ -1 \end{array} \right\} & \text{if } (i,j,k) \text{ is an } \left\{ \begin{array}{c} \text{even} \\ \text{odd} \end{array} \right\} \text{ permutation of } (1,2,3), \\ = 0 & \text{if any of } i,j,k \text{ are equal.} \end{cases} \tag{1.102}$$

An explicit calculation shows that K's and L's satisfy the following commutation relations:

$$\begin{aligned} [K_i, K_j] &= -\epsilon_{ijk} L_k \\ [L_i\,, L_j\,] &= \epsilon_{ijk} L_k \\ [L_i\,, K_j] &= \epsilon_{ijk} K_k\,, \end{aligned} \tag{1.103}$$

where sum over $k = 1, 2, 3$ is implied, and the *commutator* of two operators A, B is defined as

$$[A, B] \stackrel{\text{def}}{\equiv} AB - BA\,. \tag{1.104}$$

Note that the relation $[K_i, K_j] = -\epsilon_{ijk} L_k$ can also be written as $[K_i, K_j] = -L_k$ (i, j, k: cyclic), etc.

Exercise 1.3 *Verify the commutation relations (1.103). You may numerically verify them, or you may try proving generally by using the general formula for the elements of the matrixes.*

Exercise 1.4 *Boost in a general direction.*
Start from the formula for boost (1.9) where $P_\parallel$ is the component of $\vec{P}$ parallel to $\vec{\beta}$, and $\vec{P}_\perp$ is the component perpendicular to $\vec{\beta}$; namely,

$$P_\parallel = \vec{P} \cdot \vec{n}, \qquad \text{and} \quad \vec{P}_\perp = \vec{P} - P_\parallel \vec{n}$$

with $\vec{n} = \vec{\beta}/\beta$ (and similarly for $\vec{P}'$). Note that $\vec{\beta}$ is well-defined in the primed frame also by the particular relative orientation of the two frames chosen.
(a) Show that the corresponding Lorentz transformation matrix is given by

$$\Lambda = \begin{pmatrix} \gamma & \gamma\beta_x & \gamma\beta_y & \gamma\beta_z \\ \gamma\beta_x & 1 + \rho\beta_x^2 & \rho\beta_x\beta_y & \rho\beta_x\beta_z \\ \gamma\beta_y & \rho\beta_x\beta_y & 1 + \rho\beta_y^2 & \rho\beta_y\beta_z \\ \gamma\beta_z & \rho\beta_x\beta_z & \rho\beta_y\beta_z & 1 + \rho\beta_z^2 \end{pmatrix}, \quad \text{with} \quad \rho \equiv \frac{\gamma - 1}{\beta^2}.$$

(b) Show that when β is small, the Lorentz transformation matrix for a boost is given to the first order in β by

$$\Lambda = 1 + \beta_i K_i\,. \quad (\textit{summed over } i = 1, 2, 3)$$

(c) In the explicit expression of Λ given above, one notes that the top row $[\Lambda^0{}_\mu$ ($\mu = 0, 1, 2, 3$)$]$ and the left-most column $[\Lambda^\mu{}_0$ ($\mu = 0, 1, 2, 3$)$]$ are nothing but the velocity 4-vector $\eta^\mu = (\gamma, \vec{\beta}\gamma)$. Let's see how it works for general Lorentz transformations (proper and orthochronous). Suppose the relative orientation of the two frames K and K' is not given by $\vec{\beta}'_K = -\vec{\beta}_{K'}$, where $\vec{\beta}'_K$ is the velocity of the origin of K measured in K', and $\vec{\beta}_{K'}$ is the velocity of the origin of K' measured in K. Let Λ be the corresponding Lorentz transformation. Express $\Lambda^0{}_\mu$ and $\Lambda^\mu{}_0$ in terms of $\vec{\beta}'_K$ and $\vec{\beta}_{K'}$. (hint: Place a mass m at the origin of K and view it from K', and place a mass at the origin of K' and view it from K.)

1.7.2 Finite transformations

Now we will show that any finite (namely, not infinitesimal) rotation can be written as $e^{\theta_i L_i}$, and any finite boost can be written as $e^{\xi_i K_i}$, where θ_i and ξ_i ($i = 1, 2, 3$) are some finite real numbers. First, however, let us review some mathematical background:

Matrix exponentiations
The exponential of a $m \times m$ matrix A is also a $m \times m$ matrix defined by

$$\boxed{e^A \stackrel{\text{def}}{\equiv} \lim_{n\to\infty} \left(I + \frac{A}{n}\right)^n}, \tag{1.105}$$

which can be expanded on the right hand side as

$$e^A = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{n(n-1)\dots(n-k+1)}{k\,!} \frac{A^k}{n^k}\,. \tag{1.106}$$

Since the sum is a rapidly converging series, one can sum only the terms with $k \ll n$ for which $n(n-1)\dots(n-k+1) \approx n^k$. It then leads to

$$\boxed{e^A = \sum_{k=0}^{\infty} \frac{A^k}{k\,!}}, \tag{1.107}$$

which can also be regarded as a definition of e^A.

Using the definition (1.105) or (1.107), we see that

$$\boxed{\left(e^A\right)^\dagger = e^{A^\dagger}}, \tag{1.108}$$

where the hermitian conjugate of a matrix A is defined by $(A^\dagger)_{ij} \equiv A^*_{ji}$. The determinant of e^A can be written using (1.105) as

$$\begin{aligned} \det e^A &= \lim_{n\to\infty} \left[\det\left(I + \frac{A}{n}\right)\right]^n \\ &= \lim_{n\to\infty} \left(1 + \frac{\mathrm{Tr}A}{n} + \dots + \frac{c_k}{n^k} + \dots\right)^n, \end{aligned} \tag{1.109}$$

where we have used (1.56). This does not depend on c_k ($k > 1$) since the derivative with respect to c_k vanishes as can be readily verified. Thus, c_k ($k > 1$) can be set to zero and we have

$$\boxed{\det e^A = e^{\mathrm{Tr}A}}. \tag{1.110}$$

The derivative of e^{xA} (x is a number, while A is a constant matrix) with respect to x can be obtained using (1.107),

$$\frac{d}{dx}e^{xA} = \sum_{k=1}^{\infty} \frac{(k\,x^{k-1})A^k}{k\,!} = A\sum_{k=1}^{\infty} \frac{x^{k-1}A^{k-1}}{(k-1)\,!}\,; \tag{1.111}$$

thus,

$$\boxed{\frac{d}{dx}e^{xA} = Ae^{xA}}\,. \tag{1.112}$$

There is an important theorem that expresses a product of two exponentials in terms of single exponential, called the *Campbell-Baker-Hausdorff (CBH) theorem* (presented here without proof):

$$\boxed{e^A e^B = e^{A+B+\frac{1}{2}[A,B]+\cdots}}\,, \tag{1.113}$$

where '$\cdots$' denotes the higher-order commutators of A and B such as $[A,[A,B]]$, $[A,[[A,B],B]]$ etc. with *known coefficients*. Note that the innermost commutator is always $[A,B]$ since otherwise it is zero ($[A,A]=[B,B]=0$), and thus if $[A,B]$ is a commuting quantity (a c-number), then '$\cdots$' is zero. Applying (1.113) to $B=-A$, we get

$$e^A e^{-A} = e^{A-A} = I\,, \tag{1.114}$$

or

$$\boxed{(e^A)^{-1} = e^{-A}}\,. \tag{1.115}$$

$\blacksquare$

Rotation

An infinitesimal rotation around the z-axis by $\delta\theta$ [Figure 1.3(a)] can be written as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x - \delta\theta y \\ y + \delta\theta x \end{pmatrix} = \begin{pmatrix} 1 & -\delta\theta \\ \delta\theta & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (I + \delta\theta L_z)\begin{pmatrix} x \\ y \end{pmatrix}\,, \tag{1.116}$$

with

$$L_z = \begin{array}{c} \\ x \\ y \end{array} \!\!\begin{array}{c} \begin{array}{cc} x & y \end{array} \\ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{array}. \tag{1.117}$$

Then, a rotation by a finite angle θ is constructed as n consecutive rotations by θ/n each and taking the limit $n\to\infty$. Using (1.116), it can be written as

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= \lim_{n\to\infty}\left(I + \frac{\theta}{n}L_z\right)^n \begin{pmatrix} x \\ y \end{pmatrix} \\ &= e^{\theta L_z}\begin{pmatrix} x \\ y \end{pmatrix}\,, \end{aligned} \tag{1.118}$$

where we have used the definition (1.105).

From the explicit expression of L_z (1.117), we have $L_z^2 = -I$, $L_z^3 = -L_z$, $L_z^4 = I$, etc. In general,

$$L_z^{4n} = I, \quad L_z^{4n+1} = L_z, \quad L_z^{4n+2} = -I, \quad L_z^{4n+3} = -L_z, \tag{1.119}$$

where n is an integer. Using the second definition of e^A (1.107), the rotation matrix $e^{\theta L_z}$ can then be written in terms of the trigonometirc functions as

$$\begin{aligned} e^{\theta L_z} &= I + \theta L_z + \frac{\theta^2}{2!}\underbrace{L_z^2}_{-I} + \frac{\theta^3}{3!}\underbrace{L_z^3}_{-L_z} + \dots & (1.120)\\ &= \underbrace{\left(1 - \frac{\theta^2}{2!} + \dots\right)}_{\cos\theta} I + \underbrace{\left(\theta - \frac{\theta^3}{3!} + \dots\right)}_{\sin\theta} L_z & (1.121)\\ &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, & (1.122)\end{aligned}$$

which is probably a more familiar form of a rotation around the z-axis by an angle θ.

Similarly, rotations around x and y axes are generated by L_x and L_y as obtained by cyclic permutations of (x, y, z) in the derivation above. Switching to numerical indices $[(L_x, L_y, L_z) \equiv (L_1, L_2, L_3)]$,

$$L_1 = \begin{array}{c c} & \begin{array}{cc} 2 & 3 \end{array} \\ \begin{array}{c} 2 \\ 3 \end{array} & \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{array}, \quad L_2 = \begin{array}{c c} & \begin{array}{cc} 3 & 1 \end{array} \\ \begin{array}{c} 3 \\ 1 \end{array} & \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{array}, \quad L_3 = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ 2 \end{array} & \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \end{array}. \tag{1.123}$$

Are these identical to the definition (1.100) which was given in 4×4 matrix form, or equivalently (1.101)? Since $\delta\theta L_i$ is the *change* of coordinates by the rotation, the

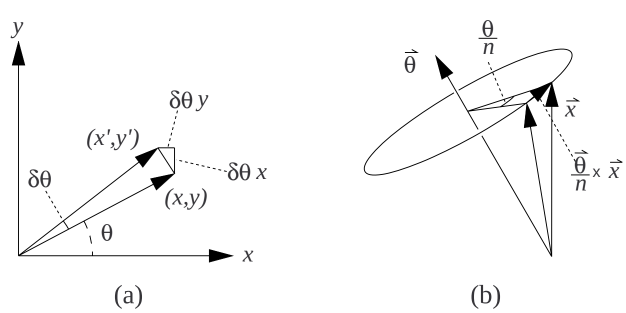

Figure 1.3: Infinitesimal rotation around the z-axis by an angle $\delta\theta$ (a), and around a general direction $\vec{\theta}$ by an angle θ/n (b).

elements of a 4×4 matrix corresponding to unchanged coordinates should be zero. We then see that the L's given above are indeed identical to (1.100).

A general rotation is then given by

$$e^{\theta_i L_i} = e^{\vec{\theta}\cdot\vec{L}}\,, \tag{1.124}$$

where

$$\vec{\theta} \stackrel{\text{def}}{\equiv} (\theta_1, \theta_2, \theta_3), \quad \vec{L} \stackrel{\text{def}}{\equiv} (L_1, L_2, L_3)\,. \tag{1.125}$$

As we will see below, this is a rotation around the direction $\vec{\theta}$ by an angle $\theta \equiv |\vec{\theta}|$. To see this, first we write $e^{\theta_i L_i}$ using the definition (1.105):

$$e^{\theta_i L_i} = \lim_{n\to\infty}\left(I + \frac{\theta_i L_i}{n}\right)^n\,, \tag{1.126}$$

which shows that it is a series of small rotations each given by $I + \theta_i L_i/n$. The action of such an infinitesimal transformation [Figure 1.3(b)] on $\vec{x}$ is (writing the space components only)

$$\begin{aligned} x'^j &= \left(I + \frac{\theta_i L_i}{n}\right)^j{}_k x^k \\ &= g^j{}_k x^k + \frac{1}{n}\theta_i \underbrace{(L_i)^j{}_k}_{-\epsilon_{ijk} \text{ by (1.101)}} x^k \\ &= x^j - \frac{1}{n}\epsilon_{ijk}\,\theta_i x^k \\ &= x^j + \frac{1}{n}(\vec{\theta}\times\vec{x})^j \end{aligned} \tag{1.127}$$

where we have used the definition of the three-dimensional cross product

$$(\vec{a}\times\vec{b})_i = \epsilon_{ijk}\, a_j b_k\,. \tag{1.128}$$

Thus, $I + \theta_i L_i/n$ is nothing but a small rotation around $\vec{\theta}$ by an angle θ/n (Figure 1.3). Then n such rotations applied successively will result in a rotation by an angle θ around the same axis $\vec{\theta}$.

Boosts

A boost in x direction by a velocity β is given by (1.26):

$$\Lambda = \begin{array}{c c} & \begin{array}{cc} t & x \end{array} \\ \begin{array}{c} t \\ x \end{array} & \begin{pmatrix} \gamma & \eta \\ \eta & \gamma \end{pmatrix} \end{array}, \qquad \left(\gamma = \frac{1}{\sqrt{1-\beta^2}}, \quad \eta = \beta\gamma\right)\,. \tag{1.129}$$

When β is small ($=\delta$), $\gamma \approx 1$ and $\eta \approx \delta$ to the first order in δ; then, the infinitesimal boost can be written as

$$\Lambda = \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix} = I + \delta K_x \,, \tag{1.130}$$

with

$$K_x = \begin{array}{c c} & \begin{array}{cc} t & x \end{array} \\ \begin{array}{c} t \\ x \end{array} & \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array} . \tag{1.131}$$

Suppose we apply n such boosts consecutively, where we take n to infinity while $n\delta$ is fixed to a certain value ξ:

$$n\delta = \xi \,. \tag{1.132}$$

Then the resulting transformation is

$$\Lambda = \lim_{n\to\infty} \left(I + \frac{\xi}{n} K_x \right)^n = e^{\xi K_x} \,, \tag{1.133}$$

where we have used the definition (1.105). Is ξ the velocity of this boost? The answer is no, even though it is a function of the velocity. Let's expand the exponential above by the second definition (1.105) and use $K_x^2 = I$:

$$\begin{aligned} \Lambda &= e^{\xi K_x} && (1.134) \\ &= I + \xi K_x + \frac{\xi^2}{2\,!} \underbrace{K_x^2}_{I} + \frac{\xi^3}{3\,!} \underbrace{K_x^3}_{K_x} + \ldots \\ &= \underbrace{\left(1 + \frac{\xi^2}{2\,!} + \ldots \right)}_{\cosh \xi} I + \underbrace{\left(\xi + \frac{\xi^2}{2\,!} + \ldots \right)}_{\sinh \xi} K_x \\ &= \begin{pmatrix} \cosh \xi & \sinh \xi \\ \sinh \xi & \cosh \xi \end{pmatrix} . && (1.135) \end{aligned}$$

Comparing with (1.130), we see that this is a boost of a velocity β given by

$$\gamma = \cosh \xi, \qquad \eta = \sinh \xi \tag{1.136}$$

or

$$\beta = \frac{\eta}{\gamma} = \tanh \xi \,. \tag{1.137}$$

Note that the relation $\gamma^2 - \eta^2 = 1$ (1.6) is automatically satisfied since $\cosh \xi^2 - \sinh \xi^2 = 1$.

Thus, n consecutive boosts by a velocity ξ/n each did not result in a boost of a velocity ξ; rather, it was a boost of a velocity $\beta = \tanh \xi$. This breakdown of the simple addition rule of velocity is well known: the relativistic rule of velocity addition states that two consecutive boosts, by β_1 and by β_2, do not result in a boost of $\beta_1+\beta_2$, but in a boost of a velocity β_0 given by

$$\beta_0 = \frac{\beta_1+\beta_2}{1+\beta_1\beta_2}\,. \tag{1.138}$$

Due to the identity $\tanh(\xi_1+\xi_2) = (\tanh\xi_1+\tanh\xi_1)/(1+\tanh\xi_1\tanh\xi_2)$, however, it becomes additive when velocities are transformed by $\beta_i = \tanh\xi_i$ $(i = 0,1,2)$; namely, $\xi_0 = \xi_1+\xi_2$ holds.

The matrix $K_x(\equiv K_1)$ given in (1.131) is identical to the 4×4 form given in (1.99) when all other elements that correspond to unchanged coordinates are set to zero. The boosts along y and z directions are obtained by simply replacing x with y or z in the derivation above. Thus, we see that K_2 and K_3 given in (1.99) indeed generate boosts in y and z directions, respectively.

A boost in a general direction would then be given by

$$\Lambda = e^{\xi_i K_i}\,, \tag{1.139}$$

where $\vec\xi \equiv (\xi_1,\xi_2,\xi_3)$ are the parameters of the boost. In order to see what kind of transformation this represents, let's write it as a series of infinitesimal transformations using (1.105):

$$e^{\xi_i K_i} = \lim_{n\to\infty}\left(I+\frac{\xi_i}{n}K_i\right)^n\,. \tag{1.140}$$

From the explicit forms of K_i (1.99), we can write the infinitesimal transformation as

$$I+\frac{\xi_i}{n}K_i = I+\frac{1}{n}\left(\begin{array}{c|ccc} 0 & \xi_1 & \xi_2 & \xi_3 \\ \hline \xi_1 & & & \\ \xi_2 & & 0 & \\ \xi_3 & & & \end{array}\right)\,. \tag{1.141}$$

On the other hand, a boost in a general direction by a small velocity $\vec\delta$ is given by (1.9) with $\gamma\approx1$, $\eta\approx\delta$ and $\delta\equiv|\vec\delta|$:

$$\begin{pmatrix} E' \\ P'_\parallel \end{pmatrix} = \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix}\begin{pmatrix} E \\ P_\parallel \end{pmatrix}\,, \qquad \vec P'_\perp = \vec P_\perp\,, \tag{1.142}$$

or

$$\left\{\begin{array}{lcl} E' & = & E+\delta P_\parallel \\ P'_\parallel & = & P_\parallel+\delta E \\ \vec P'_\perp & = & \vec P_\perp \end{array}\right. \quad \begin{array}{l} \to E' = E+\vec\delta\cdot\vec P \\ \left.\begin{array}{l} \\ \\ \end{array}\right\} \to \vec P' = \vec P + E\,\vec\delta \end{array} \tag{1.143}$$

where we have used $\vec{P}^{(\prime)} = P_\parallel^{(\prime)}\hat{\delta} + \vec{P}_\perp^{(\prime)}$ $(\hat{\delta} \equiv \vec{\delta}/\delta)$. This can be written in 4×4 matrix form as

$$\begin{pmatrix} E' \\ P'_x \\ P'_y \\ P'_z \end{pmatrix} = \left[I + \left(\begin{array}{c|ccc} 0 & \delta_1 & \delta_2 & \delta_3 \\ \hline \delta_1 & & & \\ \delta_2 & & 0 & \\ \delta_3 & & & \end{array} \right) \right] \begin{pmatrix} E \\ P_x \\ P_y \\ P_z \end{pmatrix}. \tag{1.144}$$

Comparing this with (1.141), we can identify that $I + \xi_i K_i/n$ as a boost in $\vec{\xi}$ direction by a velocity ξ/n $(\xi \equiv |\vec{\xi}|)$. Then n consecutive such boosts will result in a boost in the same direction. Since the rule of addition of velocity (1.138) is valid in any direction as long as the boosts are in the same direction, the n boosts by velocity ξ/n each will result in a single boost of velocity $\beta = \tanh\xi$ as before. Thus, $e^{\xi_i K_i}$ represents a boost in $\vec{\xi}$ direction by a velocity $\beta = \tanh\xi$.

Boost + rotation

First, we show that a rotation followed by a rotation is a rotation, but a boost followed by a boost is not in general a boost. Consider a rotation $e^{\theta_i L_i}$ followed by another rotation $e^{\phi_i L_i}$ where $\vec{\theta}$ and $\vec{\phi}$ are arbitrary vectors. Using the CBH theorem (1.113), we can write the product of the two transformations as

$$e^{\phi_i L_i} e^{\theta_j L_j} = e^{\phi_i L_i + \theta_j L_j + \frac{1}{2}[\phi_i L_i, \theta_j L_j] + \cdots}, \tag{1.145}$$

where '$\ldots$' represents terms with higher-order commutators such as $[\phi_i L_i, [\phi_j L_j, \theta_k L_k]]$ etc. Now we can use the commutation relations (1.103) to remove all commutators in the exponent on the right hand side. The result will be a linear combination of L's with well-defined coefficients (call them α_i) since the coefficients in '$\ldots$' in the CBH theorem are known. Here, there will be no K's appearing in the linear combination because of the commutation relation $[L_i, L_j] = \epsilon_{ijk} L_k$. Thus, the product is written as

$$e^{\phi_i L_i} e^{\theta_j L_j} = e^{\alpha_i L_i} \quad (\vec{\alpha} : \text{a function of } \vec{\phi}, \vec{\theta}), \tag{1.146}$$

which is just another rotation.

Next, consider a boost $e^{\xi_i K_i}$ followed by another boost $e^{\xi'_i K_i}$:

$$e^{\xi'_i K_i} e^{\xi_j K_j} = e^{\xi'_i K_i + \xi_j K_j + \frac{1}{2}[\xi'_i K_i, \xi_j K_j] + \cdots}. \tag{1.147}$$

Again, the brackets can be removed by the commutation relations (1.103) reducing the exponent to a linear combination of K's and L's. This time, there will be L's appearing through the relation $[K_i, K_j] = -\epsilon_{ijk} L_k$ which are in general not cancelled among different terms. Thus, a boost followed by a boost is not in general another boost; rather, it is a combination of boost and rotation:

$$e^{\xi'_i K_i} e^{\xi_j K_j} = e^{\alpha_i K_i + \beta_i L_i} \quad (\vec{\alpha}, \vec{\beta} : \text{functions of } \vec{\xi}, \vec{\xi'}). \tag{1.148}$$

It is easy to show, however, that if two boosts are in the same direction, then the product is also a boost.

Any combinations of boosts and rotations can then be written as

$$\boxed{\Lambda = e^{\xi_i K_i + \theta_i L_i} = e^{\frac{1}{2} a_{\mu\nu} M^{\mu\nu}}}\,, \tag{1.149}$$

where we have defined the anti-symmetric tensor $a_{\mu\nu}$ by

$$a_{0i} \stackrel{\text{def}}{\equiv} \xi_i, \quad a_{ij} \stackrel{\text{def}}{\equiv} \theta_k \ (i,j,k : \text{cyclic}), \quad a_{\mu\nu} = -a_{\nu\mu}\,, \tag{1.150}$$

and the factor $1/2$ arises since terms with $\mu > \nu$ as well as $\mu < \nu$ are included in the sum. The expression of an infinitesimal transformation (1.95) is nothing but this expression in the limit of small $a_{\mu\nu}$. Since we now know that any product of such transformations can also be written as (1.149) by the CBH theorem, we see that the set of Lorentz transformations connected to the identity is saturated by boosts and rotations.

We have seen that the generators K's and L's and their commutation relations (called the *Lie algebra*) play critical roles in understanding the Lorentz group. In fact, generators and their commutation relations completely determine the structure of the Lie group, as described briefly below.

Structure constants

When the commutators of generators of a Lie group are expressed as linear combinations of the generators themselves, the coefficients of the linear expressions are called the *structure constants* of the Lie group. For example, the coeffients $\pm\epsilon_{ijk}$ in (1.103) are the structure constants of the Lorentz group. We will now show that the structure constants completely define the structure of a Lie group. To see this, we have to define what we mean by 'same structure'. Two sets $\mathcal{F}(\ni f)$ and $\mathcal{G}(\ni g)$ are said to have the same structure if there is a mapping between $\mathcal{F}$ and $\mathcal{G}$ such that if $f_1, f_2 \in \mathcal{F}$ and $g_1, g_2 \in \mathcal{G}$ are mapped to each other:

$$f_1 \leftrightarrow g_1, \quad f_2 \leftrightarrow g_2 \tag{1.151}$$

then, the products $f_1 f_2$ and $g_1 g_1$ are also mapped to each other by the same mapping:

$$f_1 f_2 \leftrightarrow g_1 g_2\,; \tag{1.152}$$

namely, the mapping *preserves* the product rule.

Suppose the sets $\mathcal{F}$ and $\mathcal{G}$ are Lie groups with the same number of generators F_i and G_i and that they have the same set of structure constants c_{ijk}

$$[F_i, F_j] = c_{ijk} F_k, \quad [G_i, G_j] = c_{ijk} G_k\,. \tag{1.153}$$

Any element of $\mathcal{F}$ and $\mathcal{G}$ can be expressed in exponential form using the corresponding generators and a set of real parameters. We define the mapping between $\mathcal{F}$ and $\mathcal{G}$ by the same set of the real parameters:

$$f = e^{\alpha_i F_i} \leftrightarrow g = e^{\alpha_i G_i} . \tag{1.154}$$

If $f_1 \leftrightarrow g_1$ and $f_2 \leftrightarrow g_2$, then they can be written as

$$f_1 = e^{\alpha_i F_i} \;\leftrightarrow\; g_1 = e^{\alpha_i G_i} \tag{1.155}$$
$$f_2 = e^{\beta_i F_i} \;\leftrightarrow\; g_2 = e^{\beta_i G_i}, \tag{1.156}$$

where α_i and β_i are certain sets of real parameters. Then the question is whether the products $f_1 f_2$ and $g_1 g_2$ are mapped to each other by the same mapping. The products $f_1 f_2$ and $g_1 g_2$ can be written using the CBH theorem as

$$f_1 f_2 = e^{\alpha_i F_i} e^{\beta_j F_j} = e^{\alpha_i F_i + \beta_j F_j + \frac{1}{2}[\alpha_i F_i, \beta_j F_j] + \cdots} = e^{\phi_i F_i}, \tag{1.157}$$
$$g_1 g_2 = e^{\alpha_i G_i} e^{\beta_j G_j} = e^{\alpha_i G_i + \beta_j G_j + \frac{1}{2}[\alpha_i G_i, \beta_j G_j] + \cdots} = e^{\gamma_i G_i}, \tag{1.158}$$

The numbers ϕ_i and γ_i are obtained by removing the commutators using the commutation relations (1.153), and thus completely determined by α_i, β_i and c_{ijk}; namely, $\phi_i = \gamma_i$, and thus $f_1 f_2$ and $g_1 g_2$ are mapped to each other by (1.154). Thus, if two Lie groups have the same set of structure constants, then they have the same structure. ▮